AN INFORMATION-THEORETIC ANALYSIS OF GROVER'S ALGORITHM

Erdal Arikan
Electrical-Electronics Engineering Department,
Bilkent University, 06533 Ankara, Turkey
arikan@ee.bilkent.edu.tr

Abstract Grover discovered a quantum algorithm for identifying a target element in an unstructured search universe of N items in approximately $\pi/4\sqrt{N}$ queries to a quantum oracle. For classical search using a classical oracle, the search complexity is of order $N/2$ queries since on average half of the items must be searched. In work preceding Grover's, Bennett et al. had shown that no quantum algorithm can solve the search problem in fewer than $O(\sqrt{N})$ queries. Thus, Grover's algorithm has optimal order of complexity. Here, we present an information-theoretic analysis of Grover's algorithm and show that the square-root speed-up by Grover's algorithm is the best possible by any algorithm using the same quantum oracle.

Keywords: Grover's algorithm, quantum search, entropy.

1. Introduction

Grover [1], [2] discovered a quantum algorithm for identifying a target element in an unstructured search universe of N items in approximately $\pi/4\sqrt{N}$ queries to a quantum oracle oracle. For classical search using a classical oracle, the search complexity is clearly of order $N/2$ queries since on average half of the items must be searched. It has been proven that this square-root speed-up is the best attainable performance gain by any quantum algorithm. In work preceding Grover's, Bennett et al. [4] had shown that no quantum algorithm can solve the search problem in fewer than $O(\sqrt{N})$ queries. Following Grover's work, Boyer et al. [5] showed that Grover's algorithm is optimal asymptotically, and that square-root speed-up cannot be improved even if one allows, e.g., a 50% probability of error. Zalka [3] strengthened these results to show that Grover's algorithm is optimal exactly (not only asymptotically). In this correspondence we present an information-theoretic analysis of Grover's algorithm and show the optimality of Grover's algorithm from a different point of view.

A.S. Shumovsky and V.I. Rupasov (eds.), Quantum Communication and Information Technologies, 339–347.

2. A General Framework for Quantum Search

We consider the following general framework for quantum search algorithms. We let X denote the state of the target and Y the output of the search algorithm. We assume that X is uniformly distributed over the integers 0 through $N-1$. Y is also a random variable distributed over the same set of integers. The event $Y = X$ signifies that the algorithm correctly identifies the target. The probability of error for the algorithm is defined as $P_e = P(Y \neq X)$.

The state of the target is given by the density matrix density matrix

$$\rho_T = \sum_{x=0}^{N-1} (1/N)|x\rangle\langle x|, \tag{1}$$

where $\{|x\rangle\}$ is an orthonormal set. We assume that this state is accessible to the search algorithm only through calls to an oracle oracle whose exact specification will be given later. The algorithm output Y is obtained by a measurement performed on the state of the quantum computer at the end of the algorithm. We shall denote the state of the computer at time $k = 0, 1, \ldots$ by the density matrix $\rho_C(k)$. We assume that the computation begins at time 0 with the state of the computer given by an initial state $\rho_C(0)$ independent of the target state. The computer state evolves to a state of the form

$$\rho_C(k) = \sum_{x=0}^{N-1} (1/N)\rho_x(k) \tag{2}$$

at time k, under the control of the algorithm. Here, $\rho_x(k)$ is the state of the computer at time k, conditional on the target value being x. The joint state of the target and the computer at time k is given by

$$\rho_{TC}(k) = \sum_{x=0}^{N-1} (1/N)|x\rangle\langle x| \otimes \rho_x(k). \tag{3}$$

The target state (1) and the computer state (2) can be obtained as partial traces of this joint state.

We assume that the search algorithm consists of the application of a sequence of unitary operators on the joint state. Each operator takes one time unit to complete. The computation starts at time 0 and terminates at a predetermined time K, when a measurement is taken on $\rho_C(K)$ and Y is obtained. In accordance with these assumptions, we shall assume that the time index k is an integer in the range 0 to K, unless otherwise specified.

There are two types of unitary operators that may be applied to the joint state by a search algorithm: oracle oracle and non-oracle. A non-oracle operator is

of the form $I \otimes U$ and acts on the joint state as

$$\rho_{TC}(k+1) = (I \otimes U)\, \rho_{TC}(k)\, (I \otimes U)^\dagger = \sum_x (1/N)|x\rangle\langle x| \otimes U\rho_x(k)U^\dagger. \tag{4}$$

Under such an operation the computer state is transformed as

$$\rho_C(k+1) = U\rho_C(k)U^\dagger. \tag{5}$$

Thus, non-oracle operators act on the conditional states $\rho_x(k)$ uniformly; $\rho_x(k+1) = U\rho_x(k)U^\dagger$. Only oracle oracle operators have the capability of acting on conditional states non-uniformly.

An oracle operator is of the form $\sum_x |x\rangle\langle x| \otimes O_x$ and takes the joint state $\rho_{TC}(k)$ to

$$\rho_{TC}(k+1) = \sum_x (1/N)|x\rangle\langle x| \otimes O_x\rho_x(k)O_x^\dagger. \tag{6}$$

The action on the computer state is

$$\rho_C(k+1) = \sum_x (1/N)O_x\rho_x(k)O_x^\dagger. \tag{7}$$

All operators, involving an oracle or not, preserve the entropy entropy of the joint state $\rho_{TC}(k)$. The von Neumann entropy Von Neumann entropy of the joint state remains fixed at $S[\rho_{TC}(k)] = \log N$ throughout the algorithm. Non-oracle operators preserve also the entropy of the computer state; the action (5) is reversible, hence $S[\rho_C(k+1)] = S[\rho_C(k)]$. Oracle action on the computer state (7), however, does not preserve entropy; $S[\rho_C(k+1)] \neq S[\rho_C(k)]$, in general.

Progress towards identifying the target is made only by oracle oracle calls that have the capability of transferring information from the target state to the computer state. We illustrate this information transfer in the next section.

3. Grover's Algorithm

Grover's algorithm can be described within the above framework as follows. The initial state of the quantum computer is set to

$$\rho_C(0) = |s\rangle\langle s| \tag{8}$$

where

$$|s\rangle = \sum_{x=0}^{N-1} (1/\sqrt{N})|x\rangle. \tag{9}$$

Since the initial state is pure, the conditional states $\rho_x(k)$ will also be pure for all $k \geq 1$.

Grover's algorithm uses two operators: an oracle operator with

$$O_x = I - 2|x\rangle\langle x|, \tag{10}$$

and a non-oracle operator (called 'inversion about the mean') given by $I \otimes U_s$ where

$$U_s = 2|s\rangle\langle s| - I. \tag{11}$$

Both operators are Hermitian.

Grover's algorithm interlaces oracle calls with inversion-about-the-mean operations. So, it is convenient to combine these two operations in a single operation, called Grover iteration, by defining $G_x = U_s O_x$. The Grover iteration takes the joint state $\rho_{TC}(k)$ to

$$\rho_{TC}(k+1) = \sum_x (1/N)|x\rangle\langle x| \otimes G_x \rho_x(k) G_x^\dagger \tag{12}$$

In writing this, we assumed, for notational simplicity, that G_x takes one time unit to complete, although it consists of the succession of two unit-time operators.

Grover's algorithm consists of $K = (\pi/4)\sqrt{N}$ successive applications of Grover's iteration beginning with the initial state (8), followed by a measurement on $\rho_C(K)$ to obtain Y. The algorithm works because the operator G_x can be interpreted as a rotation of the x–s plane by an angle $\theta = \arccos(1-2/N) \approx 2/\sqrt{N}$ radians. So, in K iterations, the initial vector $|s\rangle$, which is almost orthogonal to $|x\rangle$, is brought into alignment with $|x\rangle$.

Grover's algorithm lends itself to exact calculation of the eigenvalues of $\rho_C(k)$, hence to computation of its entropy. The eigenvalues of $\rho_C(k)$ are

$$\lambda_1(k) = \cos^2(\theta k) \tag{13}$$

of multiplicity 1, and

$$\lambda_2(k) = \frac{\sin^2(\theta k)}{N-1} \tag{14}$$

of multiplicity $N-1$. The entropy of $\rho_C(k)$ is given by

$$S(\rho_C(k)) = -\lambda_1(k)\log\lambda_1(k) - (N-1)\lambda_2(k)\log\lambda_2(k) \tag{15}$$

and is plotted in Fig. 1 for $N = 2^{20}$. (Throughout the paper, the unit of entropy is bits and $\log$ denotes base 2 logarithm.) The entropy $S(\rho_c(k))$ has period $\pi/\theta \approx (\pi/2)\sqrt{N}$.

Our main result is the following lower bound on time-complexity.

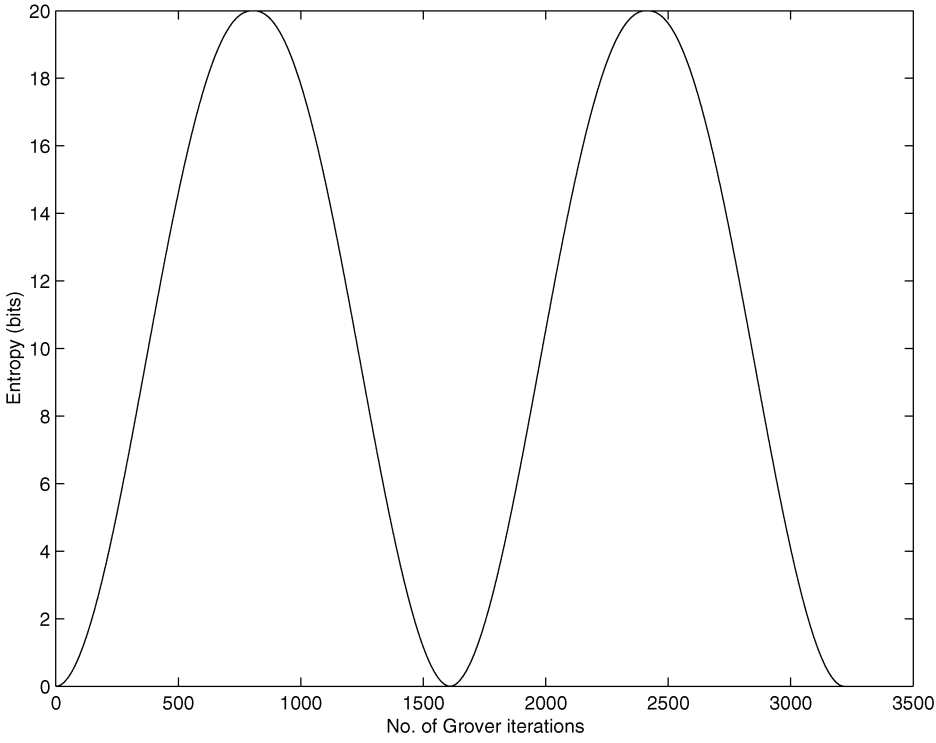

Figure 1. Evolution of entropy in Grover's algorithm.

Proposition 1 *Any quantum search algorithm that uses the oracle calls* $\{O_x\}$ *as defined by (10) must call the oracle at least*

$$K \geq \left(\frac{1 - P_e}{2\pi} + \frac{1}{\pi \log N} \right) \sqrt{N} \tag{16}$$

times to achieve a probability of error P_e.

For the proof we first derive an information-theoretic inequality. For any quantum search algorithm of the type described in section 2, we have by Fano's inequality,

$$H(Y|X) \leq \mathcal{H}(P_e) + P_e \log(N-1) \leq \mathcal{H}(P_e) + P_e \log(N), \tag{17}$$

where for any $0 \leq u \leq 1$

$$\mathcal{H}(u) = -\delta \log \delta - (1-\delta) \log(1-\delta). \tag{18}$$

On the other hand,

$$\begin{aligned} H(X|Y) &= H(X) - I(X;Y) \\ &= \log N - I(X;Y) \\ &\geq \log N - S(\rho_C(K)) \end{aligned} \tag{19}$$

where in the last line we used Holevo's bound [6, p. 531].

Let μ_k be the largest eigenvalue (sup-norm) of $\rho_C(k)$. We observe that μ_k begins at time 0 with the value 1 and evolves to the final value μ_K at the termination of the algorithm. We have

$$\begin{gathered} S(\rho_C(K)) \\ -\mu_K \log \mu_K - (1-\mu_K) \log[(1-\mu_K)/(N-1)] \end{gathered} \tag{20}$$

$$\mathcal{H}(\mu_K) + (1-\mu_K) \log N. \tag{21}$$

since the entropy entropy is maximized, for a fixed μ_K, by setting the remaining $N-1$ eigenvalues equal to $(1-\mu_K)/(N-1)$. Combining (19) and (21),

$$\mu_K \log N \leq P_e \log N + \mathcal{H}(P_e) + \mathcal{H}(\mu_K) \leq P_e \log N + 2 \tag{22}$$

Now, let

$$\Delta = \sup\{|\mu_{k+1} - \mu_k| : \ k = 0, 1, \ldots, K-1\}. \tag{23}$$

This is the maximum change in the sup norm of $\rho_C(k)$ per algorithmic step. Clearly,

$$K \geq \frac{1-\mu_K}{\Delta}.$$

Using the inequality (22), we obtain

$$K \geq \frac{1 - P_e + 2/\log N}{\Delta}. \tag{24}$$

Thus, any upper bound on Δ yields a lower bound on K. The proof will be completed by proving

Lemma 1 $\Delta \leq 2\pi/\sqrt{N}$.

We know that operators that do not involve oracle calls do not change the eigenvalues, hence the sup norm, of $\rho_C(k)$. So, we should only be interested in bounding the perturbation of the eigenvalues of $\rho_C(k)$ as a result of an oracle call. We confine our analysis to the oracle operator (10) that the Grover algorithm uses.

For purposes of this analysis, we shall consider a continuous-time representation for the operator O_x so that we may break the action of O_x into infinitesimal time steps. So, we define the Hamiltonian

$$H_x = -\pi|x\rangle\langle x| \tag{25}$$

and an associated evolution operator

$$O_x(\tau) = e^{-i\tau H_x} = I + (e^{i\pi\tau} - 1)|x\rangle\langle x|.$$

The operator O_x is related to $O_x(\tau)$ by $O_x = O_x(1)$.

We extend the definition of conditional density to continuous time by

$$\rho_x(k_0 + \tau) = O_x(\tau)\rho_x(k_0)O_x(\tau)^\dagger \tag{26}$$

for $0 \leq \tau \leq 1$. The computer state in continuous-time is defined as

$$\rho_C(t) = \sum_x (1/N)\rho_x(t). \tag{27}$$

Let $\{\lambda_n(t), u_n(t)\}$, $n = 1, \ldots, N$, be the eigenvalues and associated normalized eigenvectors of $\rho_C(t)$. Thus,

$$\rho_C(t)|u_n(t)\rangle = \lambda_n(t)|u_n(t)\rangle, \quad \langle u_n(t)|\rho_C(t) = \lambda_n(t)\langle u_n(t)|,$$

$$\langle u_n(t)|u_m(t)\rangle = \delta_{n,m}. \tag{28}$$

Since $\rho_C(t)$ evolves continuously, so do $\lambda_n(t)$ and $u_n(t)$ for each n.

Now let $(\lambda(t), u(t))$ be any one of these eigenvalue-eigenvector pairs. By a general result from linear algebra (see, e.g., Theorem 6.9.8 of Stoer and Bulirsch [7, p. 389] and the discussion on p. 391 of the same book),

$$\frac{d\lambda(t)}{dt} = \langle u(t)|\frac{d\rho_C(t)}{dt}|u(t)\rangle. \tag{29}$$

To see this, we differentiate the two sides of the identity $\lambda(t) = \langle u(t)|\rho_C(t)|u(t)\rangle$, to obtain

$$\begin{aligned}
\frac{d\lambda(t)}{dt} &= \langle u'(t)|\rho_C(t)|u(t)\rangle + \langle u(t)|\frac{d\rho_C(t)}{dt}|u(t)\rangle + \langle u(t)|\rho_C(t)|u'(t)\rangle \\
&= \langle u(t)|\frac{d\rho_C(t)}{dt}|u(t)\rangle + \lambda(t)[\langle u'(t)|u(t)\rangle + \langle u(t)|u'(t)\rangle] \\
&= \langle u(t)|\frac{d\rho_C(t)}{dt}|u(t)\rangle + \lambda(t)\frac{d}{dt}\langle u(t)|u(t)\rangle \\
&= \langle u(t)|\frac{d\rho_C(t)}{dt}|u(t)\rangle
\end{aligned}$$

where the last line follows since $\langle u(t)|u(t)\rangle \equiv 1$.

Differentiating (27), we obtain

$$\frac{d\rho_C(t)}{dt} = \sum_x -(i/N)[H_x, \rho_x(t)] \tag{30}$$

where $[\cdot,\cdot]$ is the commutation operator. Substituting this into (29), we obtain

$$\begin{aligned}
\left|\frac{d\lambda(t)}{dt}\right| &= \left|\langle u(t)| - \frac{i}{N}\sum_x [H_x, \rho_x(t)]\,|u(t)\rangle\right| \\
&\leq \frac{2}{N}\left|\sum_x \langle u(t)|H_x\rho_x(t)|u(t)\rangle\right| \\
&\overset{(a)}{\leq} \frac{2}{N}\sqrt{\sum_x \langle u(t)|H_x^2|u(t)\rangle}\ \sqrt{\sum_x \langle u(t)|\rho_x^2(t)|u(t)\rangle} \\
&\overset{(b)}{=} \frac{2}{N}\sqrt{\sum_x \pi^2\,|\langle u(t)|x\rangle|^2}\ \sqrt{N\langle u(t)|\rho_C(t)|u(t)\rangle} \\
&= \frac{2\pi}{\sqrt{N}}\,1\cdot\sqrt{\lambda(t)} \\
&\leq \frac{2\pi}{\sqrt{N}}
\end{aligned}$$

where (a) is the Cauchy-Schwarz inequality, (b) is due to (i) $\rho_x^2(t) = \rho_x(t)$ as it is a pure state, and (ii) the definition (27). Thus,

$$|\lambda(k_0+1) - \lambda(k_0)| = \left|\int_{k_0}^{k_0+1} \frac{d\lambda(t)}{dt}dt\right| \leq 2\pi/\sqrt{N}. \tag{31}$$

Since this bound is true for any eigenvalue, the change in the sup norm of $\rho_C(t)$ is also bounded by $2\pi/\sqrt{N}$.

Discussion

The bound (16) captures the $\sqrt{N}$ complexity of Grover's search algorithm. As mentioned in the Introduction, lower-bounds on Grover's algorithm have been known before; and, in fact, the present bound is not as tight as some of these earlier ones. The significance of the present bound is that it is largely based on information-theoretic concepts. Also worth noting is that the probability of error P_e appears explicitly in (16), unlike other bounds known to us.

References

[1] L. K. Grover, 'A fast quantum mechanical algorithm for database search,' *Proceedings, 28th Annual ACM Symposium on the Theory of Computing (STOC)*, May 1996, pp. 212-219. (quant-p/9605043)

[2] L. K. Grover, 'Quantum mechanics helps in searching for a needle in a haystack,' *Phys. Rev. Letters,* 78(2), 325-328, 1997. (quant-ph/9605043)

[3] C. Zalka, 'Grover's quantum searching is optimal,' *Phys. Rev. A,* 60, 2746 (1999). (quant-ph/9711070v2)

[4] C. H. Bennett, E. Bernstein, G. Brassard, and U. V. Vazirani, 'Strength and weaknesses of quantum computing,' *SIAM Journal on Computing,* vol. 26, no. 5, pp. 1510-1523, Oct. 1997. (quant-ph/9701001)

[5] M. Boyer, G. Brassard, P. Hoeyer, and A. Tapp, 'Tight bounds on quantum computing,' *Proceedings 4th Workshop on Physics and Computation,* pp. 36-43, 1996. Also *Fortsch. Phys.* 46(1998) 493-506. (quant-ph/9605034)

[6] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information,* Cambridge University Press, 2000.

[7] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis.* Springer, NY: 1980.